

Figure S1. Observed heterozygosity of SNPs ascertained from different sources. SNPs were classified into A) 180 Malawi SNPs, B) 21 Victoria SNPs, C) 9 Tanganyika SNPs, and D) 9 burtoni SNPs (see text). Boxes mark average heterozygosity, with ± 1 S.D. error bars. Generally, higher average heterozygosity is observed in those lineages where SNPs were ascertained. Heterozygosity calculated for each assemblage is generally low because these sample sets contain numerous species that are not necessarily all polymorphic. LM, Lake Malawi; LVsup, Lake Victoria superflock; LT, Lake Tanganyika.

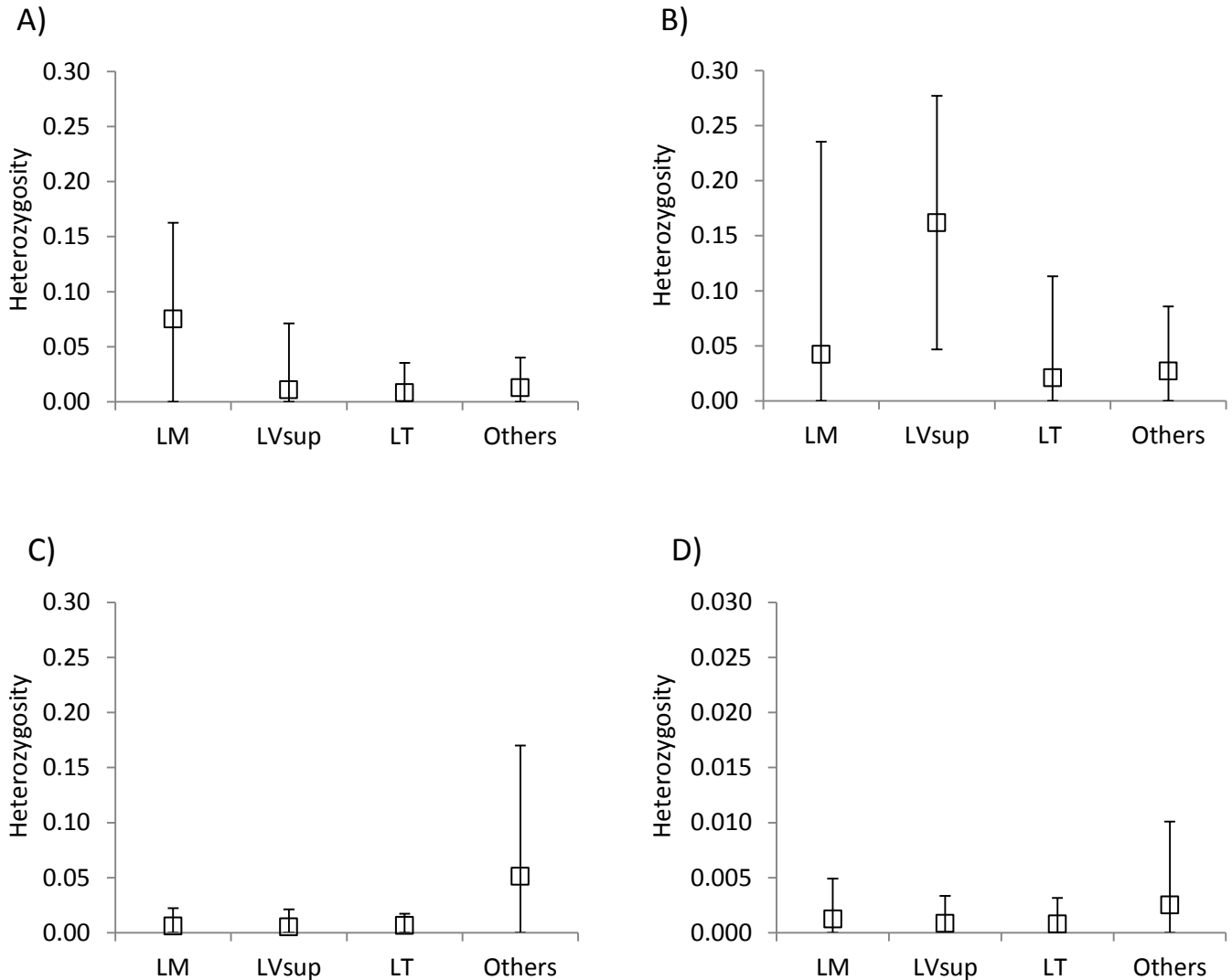


Figure S2. Cichlid population structure revealed by principal component analysis, shown plotted along the first two PC axes.

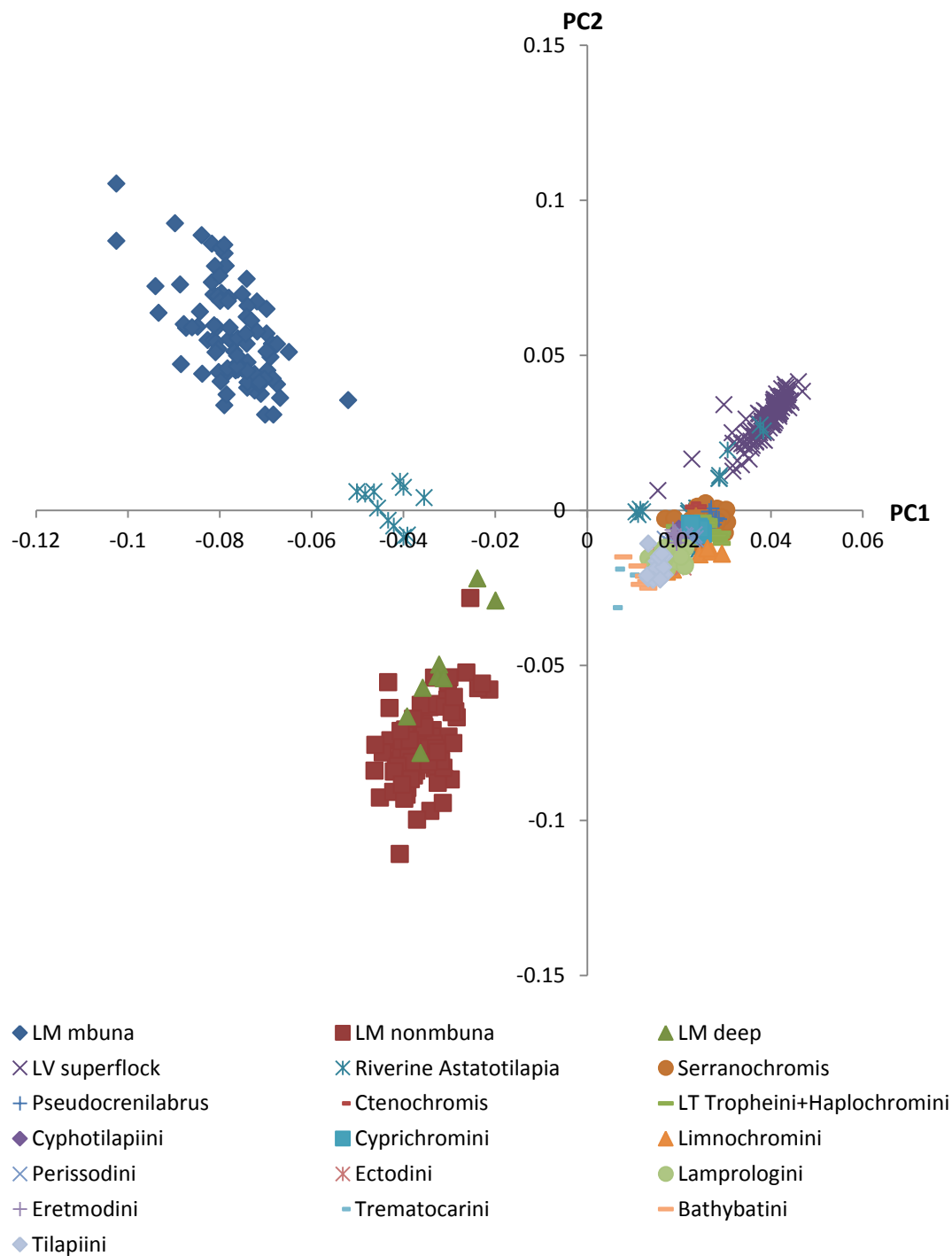


Figure S3. Consensus phylogram used in coalescent simulations to investigate causes of shared polymorphism. N_E represents the effective population sizes of the six groups, which for model simplification purposes was assumed to be equal between lineages. M is the migration rate post splitting of populations, persisting for time T_M . The star represents a population bottleneck in the Lake Victoria superflock. See text for more detail.

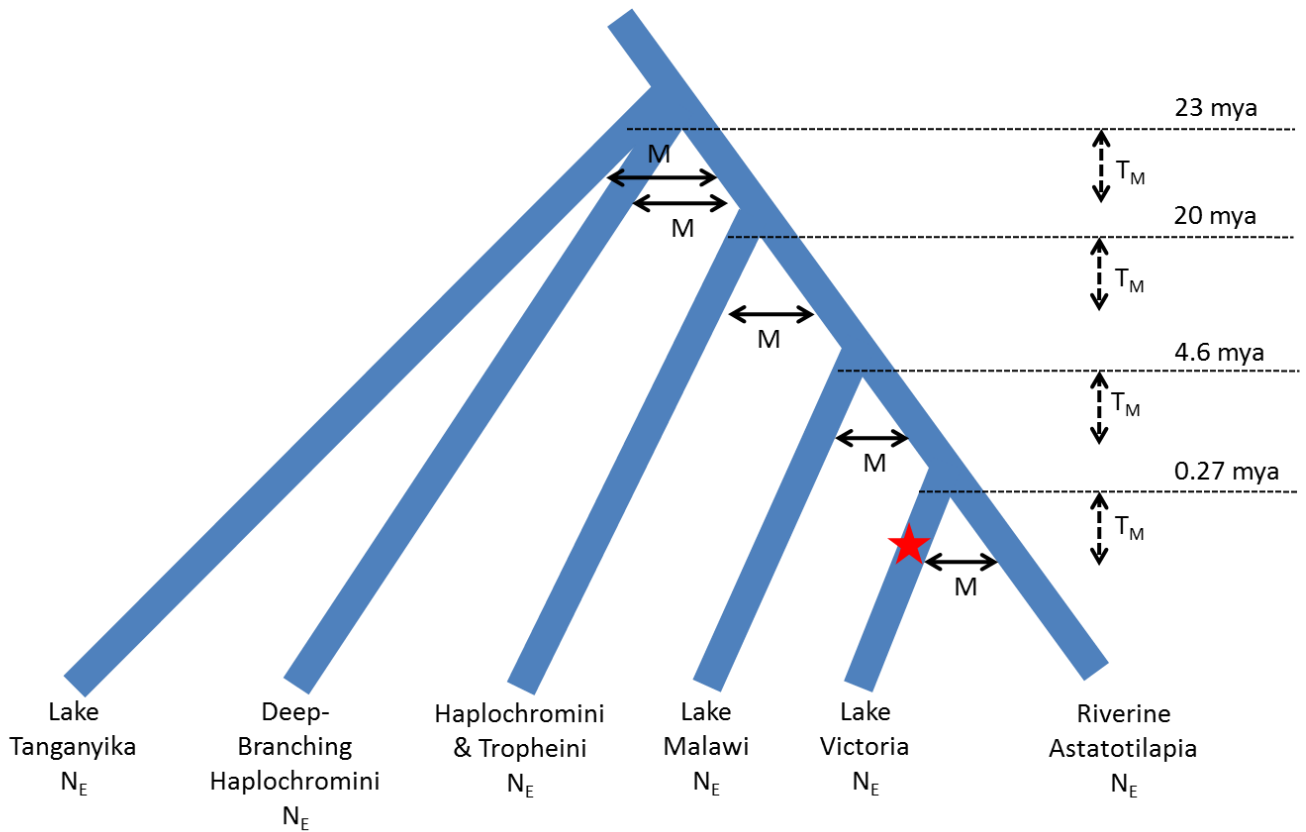


Figure S4. STRUCTURE plot from Figure 4, enlarged with sample numbers, for comparison with Table S1.

